

Light field geometry of a standard plenoptic camera

Christopher Hahne, Amar Aggoun,* Shyqyri Haxha,
Vladan Velisavljevic, and Juan Carlos Jácome Fernández

Department of Computer Science, University of Bedfordshire
Park Square, Luton, Bedfordshire, LU1 3JU, United Kingdom

[*amar.aggoun@beds.ac.uk](mailto:amar.aggoun@beds.ac.uk)

Abstract: The Standard Plenoptic Camera (SPC) is an innovation in photography, allowing for acquiring two-dimensional images focused at different depths, from a single exposure. Contrary to conventional cameras, the SPC consists of a micro lens array and a main lens projecting virtual lenses into object space. For the first time, the present research provides an approach to estimate the distance and depth of refocused images extracted from captures obtained by an SPC. Furthermore, estimates for the position and baseline of virtual lenses which correspond to an equivalent camera array are derived. On the basis of paraxial approximation, a ray tracing model employing linear equations has been developed and implemented using Matlab. The optics simulation tool Zemax is utilized for validation purposes. By designing a realistic SPC, experiments demonstrate that a predicted image refocusing distance at 3.5 m deviates by less than 11% from the simulation in Zemax, whereas baseline estimations indicate no significant difference. Applying the proposed methodology will enable an alternative to the traditional depth map acquisition by disparity analysis.

© 2014 Optical Society of America

OCIS codes: (080.3620) Lens system design; (110.5200) Photography.

References and links

1. F. Ives, "US patent 725,567," (1903).
2. G. Lippmann, "Épreuves réversibles donnant la sensation du relief," *J. Phys. Théor. Appl.* **7**, 821–825 (1908).
3. E. H. Adelson and J. R. Bergen, "The plenoptic function and the elements of early vision," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, eds. (MIT Press, 1991), pp. 3–20.
4. E. H. Adelson and J. Y. Wang, "Single lens stereo with a plenoptic camera," *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 99–106 (1992).
5. M. Levoy and P. Hanrahan, "Lightfield rendering," in *Proceedings of ACM SIGGRAPH*, (1996), pp. 31–42.
6. A. Isaksen, L. McMillan, and S. J. Gortler, "Dynamically reparameterized light fields," in *Proceedings of ACM SIGGRAPH*, (2000), pp. 297–306.
7. R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Lightfield photography with a hand-held plenoptic camera," *Stanford Tech. Report CTSR* (2005), pp. 1–11.
8. M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," in *Proceedings of ACM SIGGRAPH*, (2006), pp. 924–934.
9. M. Broxton, L. Grosenick, S. Yang, N. Cohen, A. Andalman, K. Deisseroth, and M. Levoy, "Wave optics theory and 3-D deconvolution for the light field microscope," *Opt. Express* **21**, 25418–25439 (2013).
10. A. Lumsdaine and T. Georgiev, "The focused plenoptic camera," in *Proceedings of IEEE International Conference on Computational Photography*, (IEEE, 2009), pp. 1–8.
11. T. Georgeiv, K. C. Zheng, B. Curless, D. Salesin, S. Nayar, and C. Intwala, "Spatio-angular resolution tradeoff in integral photography," in *Proceedings of Eurographics Symposium on Rendering*, (2006), pp. 263–272.
12. C. Wu, M. McCormick, A. Aggoun, and S. Y. Kung, "Depth Mapping of Integral Images Through Viewpoint Image Extraction With a Hybrid Disparity Analysis Algorithm," *Journal of Display Technology* **4**, 101–108 (2008).

13. T. E. Bishop, S. Zanetti, and P. Favaro, "Light field superresolution," in *IEEE International Conference on Computational Photography*, (2009).
14. C. Perwass and L. Wietzke, "Single-lens 3D camera with extended depth-of-field," in *Human Vision and Electronic Imaging XVII*, Proc. SPIE **8291**, 829108 (2012).
15. T. Georgiev, A. Lumsdaine, and S. Goma, "Plenoptic Principal Planes," in *Imaging and Applied Optics*, OSA Technical Digest (CD) (Optical Society of America, 2011), paper JTuD3.
16. C. Hahne and A. Aggoun "Embedded FIR filter design for real-time refocusing using a standard plenoptic video camera," in *Digital Photography X*, Proc. SPIE **9023**, 902305 (2014).
17. E. Hecht, *Optics, Fourth Edition* (Addison Wesley, 2001).
18. N. Konidakis, "Optical Prescriptions in Zemax," (2014), <https://sites.google.com/site/nickkonidakis/prescriptions>.
19. Y. Pritch, M. Ben-Ezra, and S. Peleg "Automatic disparity control in stereo panoramas (OmniStereo)," in *Proceedings of IEEE Workshop on Omnidirectional Vision*, (2000), pp. 54–61.

1. Introduction

In recent years, there has been an increasing interest in plenoptic cameras and their ability to refocus two-dimensional (2-D) images after image capture. Initial research in the subject of light field can be traced back to Ives [1] and Lippmann [2] who independently discovered the possibility of gathering light rays from different angles by an array pinholes and micro lenses, respectively. Subsequently, several studies have been produced over the last century including the seven-dimensional plenoptic function [3], the plenoptic camera [4] and the four-dimensional (4-D) light field parameterization [5]. The latter achievement simplifies the plenoptic function by describing the Light Field (LF) as a set of rays intersecting two 2-D planes enabling image acquisition and reconstruction of spatial and angular light information. In 2000, Isaksen *et al.* [6] explored the capability of refocusing using viewpoint images captured by an array of cameras. Refocusing can be seen as synthesizing a 2-D Focused Image Slice (FIS) of the 4-D LF. Ng *et al.* carried this technological idea further to investigate refocusing based on a hand-held Standard Plenoptic Camera (SPC) having a Micro Lens Array (MLA) attached in front of the sensor [7]. Afterwards, this plenoptic setup has been implemented in a microscope by Levoy *et al.* [8] and was recently advanced by Broxton *et al.* [9]. In 2009, Lumsdaine and Georgiev [10] massively improved the effective spatial resolution by introducing a new rendering technique for the Focused Plenoptic Camera (FPC) allowing for different positions of the MLA. Nevertheless, the FPC inherently causes a loss of angular information resulting in a trade-off between angular and spatial resolution [11]. An early investigation of depth measuring using a disparity analysis of integral images was conducted by Wu *et al.* [12]. The challenge with this approach is to minimize the error resulting from the disparity map which is due to the relatively small baseline compared to a camera array setup. Improvements have been made by Bishop *et al.* [13] and Perwass *et al.* [14]. First research examining the position of virtual micro lenses was undertaken by Georgiev *et al.* [15]. Therein, principal plane calculations of the FPC identified that virtual lenses are projected into object space comparable to an array of cameras. So far, the current state of that attempt does not provide baseline estimations.

In relation to the SPC, an uncertainty also exists about the virtual baseline just as the distance of an FIS. With the aid of paraxial approximation and linear equations, this paper addresses and solves these problems for the first time. The presented proposition contributes to the subject of LF in several ways. First, it assists in the specification of an SPC in advance. Secondly, the distance prediction of FISs will enhance SPC depth map computations. Moreover, since screening plenoptic content on multi-view displays requires knowledge about the baseline setup, the novel method presented in this paper supports that application. Experimental testings using the real ray tracing tool Zemax verify the Matlab implementation by a relative error below 0.5% at distances ≤ 300 mm and by less than 11% at a distance of approximately 3.5 m.

2. Ray tracing intersection model

Historically, the most influential achievement in capturing LF data was elaborated by Levoy *et al.* [5] when describing rays of an LF by intersections of two 2-D planes denoted by (u, v) and (s, t) , respectively. Subsequently, Ng *et al.* [7] reinvestigated the LF parameterization $L(s, t, u, v)$ by adapting it to a plenoptic camera in which the measured irradiance I occurring at the micro lens plane (s, t) is given by

$$I_{b_U}(s, t) = \frac{1}{b_U^2} \iint L_{b_U}(s, t, U, V) A(U, V) \cos^4 \theta \, dU \, dV \quad (1)$$

where $A(\cdot)$ denotes the aperture, (U, V) the main lens plane and b_U the spacing between the main lens and MLA plane (s, t) . The roll-off factor $\cos^4 \theta$ is also known as vignetting. By considering the aperture to be completely open, so that $A(\cdot) = 1$, and neglecting the vignetting as well as the inverse square law in terms of the factor $1/b_U^2$, the equation can be shortened. In order to further simplify following descriptions, the LF is seen to be captured by rays intersecting two one-dimensional (1-D) planes disregarding the vertical dimension. Thereby, subsequent declarations are based on the assumption that camera parameters are equally specified in horizontal and vertical dimension allowing to apply proposed solutions in both directions similarly. Hence, the irradiance is horizontally given by

$$I_{b_U}(s) = \int L_{b_U}(s, U) \, dU \quad (2)$$

On closer examination of the simplified Eq. (2) from Ng, it may be obvious that $L_{b_U}(s, U)$ is a compressed LF image projection determined by intersections at micro lens plane s and main lens plane U , compliant with the LF parameterization proposed by Levoy.

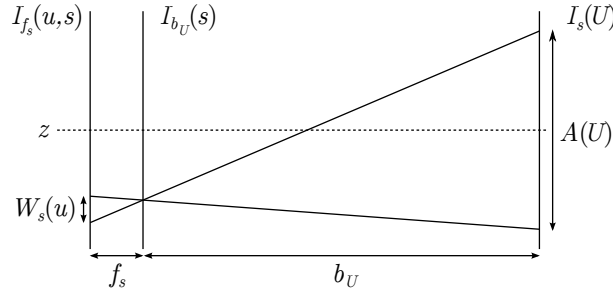


Fig. 1. Planes of irradiance (Ref. [16], Fig. 1).

However, the incident irradiance of light is actually measured at the image plane of micro lenses, denoted by u . Therefore, the two planes providing retrievable coordinates at which rays intersect are rather (u, s) than (s, U) . Neglecting reflections and absorptions due to the lens material, Fig. 1 illustrates, by means of the method of similar triangles, that the irradiance I_{b_U} at a particular point s emerging from the main lens U is proportionally distributed along its micro image $W_s(u)$ with the irradiance $I_{f_s}(u)$. So it follows that

$$I_{b_U}(s) = \int I_{f_s}(u, s) \, du \quad (3)$$

In case the wavelength of the light spectrum is limited and weighted according to the human visual perception, the irradiance can be substituted by the photometric illuminance E giving

$$E_{b_U}(s) = \int E_{f_s}(u, s) \, du \quad (4)$$

2.1. Standard plenoptic camera

A growing body of literature explains refocusing primarily by using the method of similar triangles [7, 14, 15]. According to Ng *et al.* [7], in an SPC the image plane of the sensor is placed at the distance of the micro lens focal length f_s behind the MLA. As initially shown in a research publication by Hahne and Aggoun [16], chief rays can be traced from the image plane of the sensor into the real object space by taking advantage of the thin lens equation [17]

$$\frac{1}{f_s} = \frac{1}{a_s} + \frac{1}{b_s} \quad (5)$$

Given the constraint that the image distance b_s of the micro lens equals its focal length f_s ($b_s = f_s$), it is mathematically demonstrated in

$$0 = \lim_{a_s \rightarrow \infty} \left(\frac{1}{a_s} \right) \quad (6)$$

that the basic idea of this ray tracing approach relies on the fact that in geometrical optics collimated light rays converge on the focal point of a convex lens. At this stage, it can be assumed that focused spots behind the MLA are of an infinitesimal size. As a result, chief rays impinging on infinitesimal spots at discrete positions on the image plane can be traced through the micro and main lens as illustrated in Fig. 2. Therein, positions within a micro image are described by u having a consecutive index $i \in \mathbb{Z}$ in the range of $[-(\hat{m}-1)/2, (\hat{m}-1)/2]$ where \hat{m} represents the 1-D number of pixels within each micro image that is considered to be consistent. On condition that \hat{m} is an odd number, the central micro image position can be obtained by $c = (\hat{m}-1)/2$. When starting to count the index from the central position c , micro image locations are given by u_{c+i} . The spacing of two adjacent positions u_{c+i} can be seen as the pixel pitch p_p in terms of a digital camera.

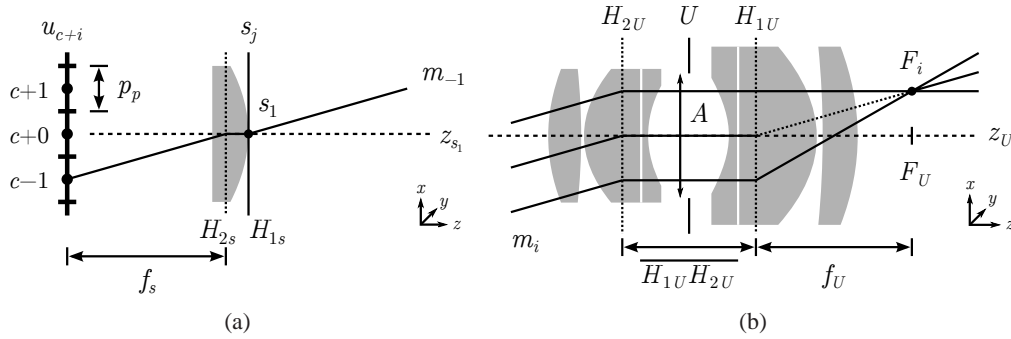
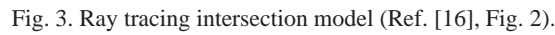


Fig. 2. (a) Micro lens s_j and a chief ray m_i . (b) Collimated light rays traveling through the main lens.

Since there is a single chief ray for each position $u + i$, chief rays may be distinguished by their respective slope m_i . For instance, in Fig. 2(a) chief rays, having the slope m_{-1} , focus at u_{c-1} under each micro lens s indexed by $j \in \mathbb{Z}$. Hence, chief rays m_{-1} under all micro lenses s form a collimated light beam m_{-1} in front of the micro lenses. Generally, the depth is represented by z and more specifically by z_{s_j} for the optical axis of the main lens and z_{s_j} for each micro lens correspondingly. As depicted in Fig. 2(b), due to the behavior of collimated light, parallel rays forming the light beam m_{c-1} are considered to be refracted at the system's principal planes H_{U1} and H_{U2} of the main lens U and diverge from the main lens focal plane F_U . This respective point is denoted by F_i and varies along the dimensions perpendicular to z .

Apart from the assumptions made earlier in this section, it is supposed that propagating light is reflected from Lambertian surfaces into the camera device with a luminous emittance M equal to the illuminance E_{b_U} . In addition, since E_{b_U} consists of spatially sampled points s_j , an object plane M can be similarly described by discrete object points s'_j . As seen in Eq. (4), synthesizing an image E'_{b_U} , which would have been captured by a conventional camera with an illuminance E_{b_U} at plane s , requires the selection and summation of specific values $E_{f_s}[u_{c+i}, s_j]$. Furthermore, a raw LF image of an SPC enables the generation of images which would have focused behind the image sensor (at distances greater than b_U) without an MLA. These hypothetic images would have an illuminance E and form the refocusable LF which is compressed by the micro lenses and distributed over $E_{f_s}(u, s)$. In order to distinguish between refocusable image planes E , E_{b_U} is substituted with a synthesized illuminance E'_a whereas $E_{b_U} = E'_0$ and image planes E'_a , being further away, are indexed consecutively by a .


$$E'_1[s_2] = E_{f_s}[u_2, s_2] + E_{f_s}[u_1, s_3] + E_{f_s}[u_0, s_4] \quad (8)$$

To make it easier to follow the idea of ray tracing in the plenoptic LF, chief rays of the given examples in Eqs. (7) and (8) are highlighted in Fig. 4. Thereby, it is apparent that object points at plane M_1 would have focused at E'_1 behind the actual image plane E_{f_s} .

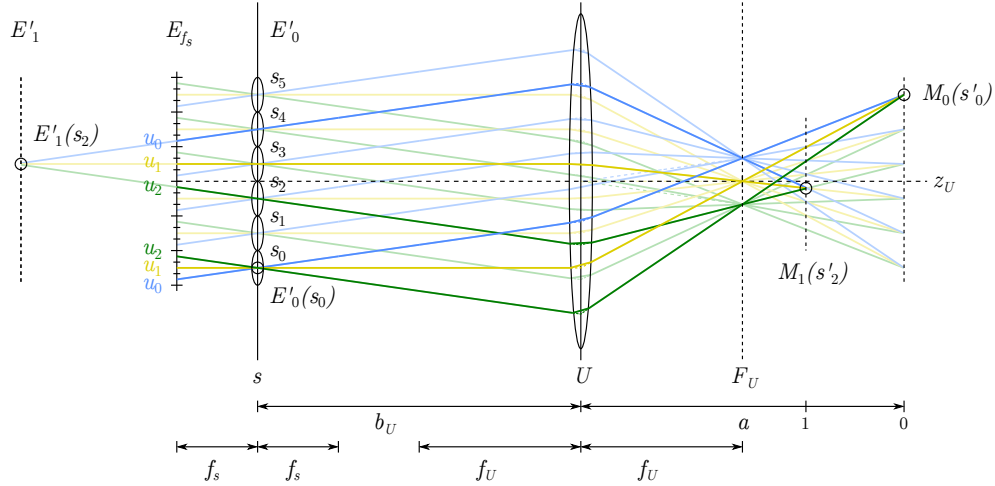


Fig. 4. Ray tracing intersection examples demonstrating Eqs. (7) and (8).

Investigating all possible synthesis combinations allows for reverse-engineering an algorithmic formula providing FISs E'_a from the raw data E_{f_s} which is given by

$$E'_a[s_j] = \sum_{i=-c}^c E_{f_s}[u_{\hat{m}-1-c+i}, s_{j+a(c-i)}], \quad a \in \mathbb{Z} \quad (9)$$

As suggested by Hahne and Aggoun [16], a translation from $[u_{c+1}, s_j]$ coordinate space to a single index $[x_k]$, as commonly designated in a digital image, is given by

$$k = j \times \hat{m} + c + i \quad (10)$$

enabling a conversion of the synthesis examples in Eqs. (7) and (8) to

$$E'_0[s_0] = E_{f_s}[x_2] + E_{f_s}[x_1] + E_{f_s}[x_0] \quad (11)$$

$$E'_1[s_2] = E_{f_s}[x_8] + E_{f_s}[x_{10}] + E_{f_s}[x_{12}] \quad (12)$$

3. Distance estimation of focused image slices

Given the background of the SPC and the refocusing synthesis, this section turns the focus on the elaboration of the FIS distance. In particular, the first subsection aims to develop a distance prediction based on chief rays at pixel centers whereas the subsequent part considers the pixel width and optical resolution limit to estimate the depth of an FIS.

3.1. Central position

Closer inspection of the syntheses of Eqs. (7) and (8) reveals that selecting only two combinations of $E_{f_s}[u_{c+i}, s_j]$ being merged and tracing the related chief rays back to the intersection in object space suffices to acquire the metric distance of a respective slice a . By subdividing the path of each chief ray into the intervals at which refractions can be seen to occur, a ray path is mathematically described as a composition of linear equations. Referring to Fig. 2, rays converging at the same relative position u_{c+i} under each micro lens s_j , have the same incidence angle, in other words the same slope m_i which is given by

$$m_i = \frac{\Delta u \times p_p}{f_s} \quad (13)$$

where $\Delta u = u_c - u_{c+i}$ and u_c denotes the position of a central focal point under a micro lens. By introducing $n_{i,j}$ to be the x -intercept at the image plane which is defined as

$$n_{i,j} = j \times p_{\hat{m}} + \frac{p_{\hat{m}}}{2} + \Delta u \times p_p \quad (14)$$

$p_{\hat{m}}$ denotes the pitch of the micro lenses s indexed by j starting from the bottom lens. A linear function $\hat{f}_{i,j}(z)$ representing a chief ray from the micro lens image plane to the optical center of the main lens U is thus formed by

$$\hat{f}_{i,j}(z) = m_i \times z + n_{i,j}, \quad z \in [0, U] \quad (15)$$

In paraxial ray tracing, a single ray can be seen to be geometrically refracted at the principal plane H of a thin lens. Depending on the slope m_i , the micro image position Δu and its parent micro lens s_j , rays intersect the optical center at point $U_{i,j}$ which is given by

$$U_{i,j} = m_i \times (f_s + b_U) + n_{i,j} \quad (16)$$

Since rays converging at Δu under s are parallel to each other in the interval of the optical centers from s to U , they also converge in front of the main lens at its focal plane F_U . In using m_i as the representation of the angle, the intersection of corresponding light rays with focal plane F_U is derived from

$$F_i = m_i \times f_U \quad (17)$$

so that each ray focusing at Δu under all s_j originates in F_i . By having calculated F_i and the individual point $U_{i,j}$, the slope $q_{i,j}$ of each light ray in object space is deduced from

$$q_{i,j} = \frac{F_i - U_{i,j}}{f_U} \quad (18)$$

In case there is no object surface at F_i where light rays can be reflected from, the corresponding ray is emanated from an object at a further distance than F_i . Given $q_{i,j}$, the path of Δu can be traced back to infinity by the function

$$\hat{f}_{i,j}(z) = q_{i,j} \times z + U_{i,j}, \quad z \in [U, \infty) \quad (19)$$

The distance from U to a point of the respective FIS, denoted by z_a , can be algebraically obtained by solving

$$\hat{f}_{-c,e}(z) = \hat{f}_{a-c,e-a}(z), \quad z \in [U, \infty) \quad (20)$$

where $\hat{f}_{-c,e}(z)$ is an exemplary reference chief ray having a slope m_{-c} since $(i, j) = (-c, e)$. Parameter e may be an arbitrary value to represent a valid micro lens s_e . To get an intersecting chief ray, the second position may be given by $(i, j) = (a - c, e - a)$. The solution z_a of Eq. (20) gives the intersection of two chief rays in object space and merely the distance from the main lens U to the corresponding FIS. When considering the impact of a thick lens or even a lens system, separations $\overline{H_{1s}H_{2s}}$ and $\overline{H_{1U}H_{2U}}$ between first and secondary principal planes are simply added to the resulting intersection distance. Note that $\overline{H_{1U}H_{2U}}$ may be negative in case principal planes are interchanged. The final distance d_a between the image plane of the sensor and the refocusing plane is therefore

$$d_a = f_s + \overline{H_{1s}H_{2s}} + b_U + \overline{H_{1U}H_{2U}} + z_a \quad (21)$$

3.2. Depth of field of a focused image slice

Due to the restriction of tracing chief rays back to the object space, only an infinitesimally thin depth range of an FIS is calculated. Contrary to the limitation on chief rays, light beams rather intersect each other in object space leading to a Depth Of Field (DOF) for each FIS. The DOF of an FIS is even larger as focused points on the sensor are not of an infinitesimal small width, but have some spacing due to the size of a pixel. Apart from the pixel pitch, optical elements of an imaging system also restrict the resolution limit $(\Delta\ell)_{min}$ by the separation of the Airy disks. In Hecht [17], it has been stated that the separation should at least equal the radius r_1 of an Airy disk center peak in order to distinguish between adjacent Airy disks. In general, this is approximately given by

$$(\Delta\ell)_{min} \approx 1.22 \frac{f\lambda}{A} \quad (22)$$

where λ denotes the wave length of light. In subsequent explanations, it is supposed that $(\Delta\ell)_{min} \leq p_p$ provided that the pixel pitch determines the resolution limit.

Assuming that locations u_{c+i} represent pixel centers, the approach described in the previous Subsection 3.1 solely reveals the central position of a respective FIS. Thus, in order to approximate the depth of an FIS, pixel borders need to be the subject of investigation. Because the illuminance samples E_{fs} are considered to have some width, DOF ray positions of a pixel $E_{fs}[u_{c+i}, s_j]$ are class-divided into three types:

- *inner rays* at pixel borders towards the micro image center
- *outer rays* at the pixel borders closer to micro image edge
- *central rays* at the pixel center

Recalling the example in Eq. (8), pixels forming $E'_1[s_2]$ are subject of investigation. Given this example, Fig. 5 demonstrates that an intersection of *inner rays*, depicted in red color, occurs at the shorter distance d_{a-} whereby intersecting *outer rays* in black color give the distance d_{a+} . Since all rays between *outer* and *inner rays* including *central* chief rays intersect at a distance in the range of d_{a-} and d_{a+} , the DOF, denoted by D_a , of an FIS at distance d_a is then

$$D_a = d_{a+} - d_{a-} \quad (23)$$

The tracing of *inner* and *outer rays* is achieved in a similar way to the *central* chief ray tracing. Taking into consideration that rays at the pixel border are separated by $p_p/2$ from the position u_{c+i} , their respective slope $m_{i\pm}$ to the optical micro lens center is

$$m_{i\pm} = \frac{\Delta u \times p_p \pm p_p/2}{f_s} \quad (24)$$

As seen in Fig. 5, *inner* and *outer rays* travel with a slope $m_{i\pm}$ from a micro lens border $s_{i,j\pm}$ which is given by

$$s_{i,j\pm} = j \times p_{\hat{m}} + p_{\hat{m}}/2 \pm p_{\hat{m}}/2 \quad (25)$$

intersecting a main lens principal plane at $U_{i,j\pm}$ which is obtained by

$$U_{i,j\pm} = m_{i\pm} \times b_U + s_{i,j\pm} \quad (26)$$

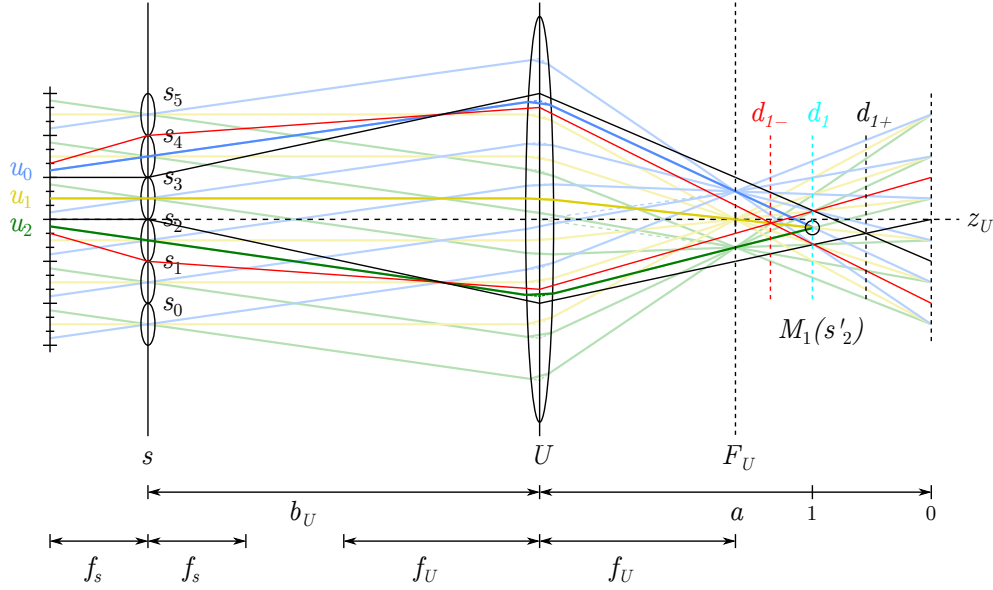


Fig. 5. Ray tracing intersection model indicating the DOF for FIS $a = 1$.

Following the suggestion made in Subsection 3.1, collimated rays reach F_U at $F_i \pm$ yielding

$$F_{i\pm} = m_{i\pm} \times f_U \quad (27)$$

Subsequently, the object space slope $q_{i,j\pm}$ is calculated from previous intersections such that

$$q_{i,j\pm} = \frac{F_{i\pm} - U_{i,j\pm}}{f_U} \quad (28)$$

As a result, a linear equation representing this ray in object space can be written in a fashion

$$\hat{f}_{i,j\pm}(z) = q_{i,j\pm} \times z + U_{i,j\pm}, \quad z \in [U, \infty) \quad (29)$$

Investigations into the example depicted in Fig. 5 show that traced light beams of most distant pixel combinations $E_{f_s}[u_{c+i}, s_j]$, which are merged in the refocusing synthesis of a particular position $E'[s_j]$, enclose the smallest area of that point in object space. In other words, an illuminance of potentially emanating light rays entering the camera device from that confined space, is covered by all pixels involved to form $E'[s_j]$. In contrast, a larger space does not fulfill this condition as the illuminance of light reflected from slightly different locations is not equally measured by all those sensor pixels. Hence, the range bordered by rays of most distant pixels ensures to obtain the maximal effective resolution of the corresponding FIS. The selection of most distant positions of integrated values $E_{f_s}[u_{c+i}, s_j]$ can be described by $(i, j) = (-c, a(\hat{m} - 1) + e)$ for the first ray position $E_{f_s}[u_0, s_4]$, where $e = 2$ according to Fig. 5, and $(i, j) = (c, e)$ for the second $E_{f_s}[u_2, s_2]$ intersecting at the distance of FIS $a = 1$. Considering this example, depth borders lie at distances $z_{1\pm}$ which are obtained by the solutions of

$$\hat{f}_{0,4\pm}(z) = \hat{f}_{2,2\pm}(z), \quad z \in [U, \infty) \quad (30)$$

for z_{1-} and z_{1+} , respectively. Similar to the central refocusing distance procedure, Eq. (30) only gives the spacing between main lens principal plane and the depth range limits. Consequently, the path from sensor to main lens has to be added in order to get $d_{a\pm}$ as given by

$$d_{a\pm} = f_s + \overline{H_{1s}H_{2s}} + b_U + \overline{H_{1U}H_{2U}} + z_{a\pm} \quad (31)$$

4. Baseline of virtual lenses

As carried out by Ng [7], rearranging each illuminance value E_{fs} , having u_{c+i} in common, to create a single image $E'_{u_{c+i}}$, provides a virtual viewpoint just as it would have been taken with a real camera array. A 1-D viewpoint image is extracted by

$$E'_{u_{c+i}}[s_j] = E[u_{c+i}, s_j] \quad (32)$$

whereas the given synthesis equation can certainly be used for the vertical dimension in the same manner. Due to that algorithm, the effective resolution of a viewpoint image acquired by an SPC corresponds to the number of micro lenses and the number of virtual lenses thus equals the micro image resolution \hat{m} .

In the absence of aberrations, the optical center of a virtual lens is optimally at the best focus of rays m_i converging at plane F_U as can be seen in the ray tracing intersection model in Fig. 3. Although, considering the pixel size, results in a larger optical center having a width w . Each virtual lens has a virtual optical axis z'_i represented by a chief ray of the respective viewpoint. As a consequence, the virtual lens plane along w and the viewpoint's virtual optical axis are at a right angle with respect to each other. Hence, when adjusting the tilt angle of a virtual lens to be parallel to the main optical axis z , the virtual optical axes of the viewpoints have to be collimated. This requirement is fulfilled when the spacing between main lens and MLA amounts to the focal length ($f_U = b_U$). The tilt angle Φ_i of a virtual lens is given by

$$\Phi_i = \arctan\left(\frac{m_i \times (b_U - f_U)}{f_U}\right) \quad (33)$$

As shown in Figs. 6(a)–6(c), the tilt angle setting of virtual lenses depends on b_U .

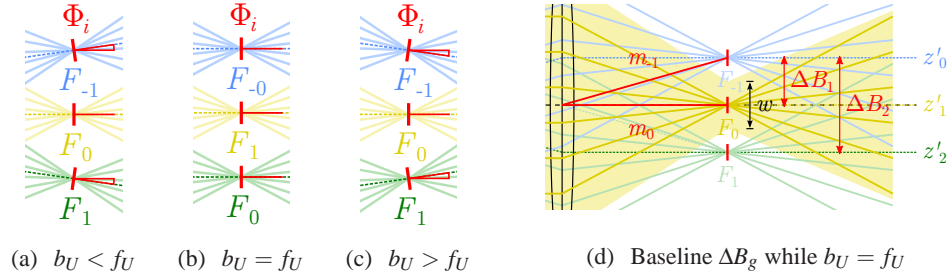


Fig. 6. (a–c) Tilt angles Φ_i of the virtual lenses F_i ; (d) Illustration of baseline ΔB_g between virtual optical axes z'_i .

When seeking F_i from Eq. (17) in Fig. 3, it may be obvious that the optical center of a virtual lens is located at position (F_U, F_i) . As a result, the horizontal positions of two virtual optical centers F_i provide the corresponding baseline ΔB_g by

$$\Delta B_g = |F_i - F_{i+g}| \quad (34)$$

where g serves as a gap between virtual lenses. Figure 6(d) illustrates different gaps g while having the main lens focused to infinity ($b_U = f_U$). Using the proposed approach, a baseline prediction of any pair of virtual lenses is possible. For example, a gap of $g = 1$ indicates direct adjacency among viewpoints. The width w of a virtual lens equals the smallest baseline ($w = \Delta B_1$) and can be visualised by tracing *inner* and *outer* rays as shown in Subsection 3.2.

5. Model validation

5.1. Implementation

In order to evaluate suggested estimations, the ray tracing attempt described in Section 3 has been implemented in Matlab. For verification purposes, the well known complex ray tracing simulation tool Zemax is used. A comparison requires the specification of a realistic SPC to be modeled equally in both environments. In Zemax, light rays have been emitted with a wavelength $\lambda = 632.8$ nm. For simplicity, the arrangement of the micro lenses is specified to be rectangular. Micro lenses are of a plano-convex shape having a focal length $f_s = 2.749$ mm and a thickness $t_s = 1.1$ mm. The MLA is made from glass substrate with a refractive index of $n_{632.8} = 1.515$. By applying

$$R_s = f_s \times (n - 1) \quad (35)$$

derived from Hecht [17], a radius of curvature $R_s \approx 1.416$ is calculated. The micro lens pitch is defined to be $p_m = 250$ μm . Modelling the specified micro lens in Zemax, it turns out that $\overline{H_{1s}H_{2s}} = 0.374$ mm. Therefore, the distance d_s from back vertex V_{s2} to image plane is given by

$$d_s = f_s - (t_s - \overline{H_{1s}H_{2s}}) \quad (36)$$

and amounts to 2.023 mm. The image sensor of the camera has a pixel pitch $p_p = 0.01$ mm which also describes the spacing between u_c and its adjacent position u_{c-1} . Referring to Eq. (13), it is apparent that the slope $m_{c-1} \approx 0.0036$ and $\arctan(m_{c-1}) \approx 0.206^\circ$ is a chief ray angle entered in Zemax. Spherical aberrations are suppressed as much as possible by choosing a Double Gauss objective which can be treated as a thick lens with two principal plane locations. Therefore, a Double Gauss 28 degree field objective, denoted by f_{100} , was taken from the Zemax sample library. Alternatively, to examine the impact of main lens parameters, a 28 – 200 mm $f/3.5 - 5.6$ zoom objective, represented by f_{200} , from Nikon is chosen which can be found online as a Zemax file containing the entire lens model [18]. Determination of principal plane positions and exact focal lengths has been accomplished by using Zemax functions. Table 1 summarizes the main lens data essential for the ray tracing approach.

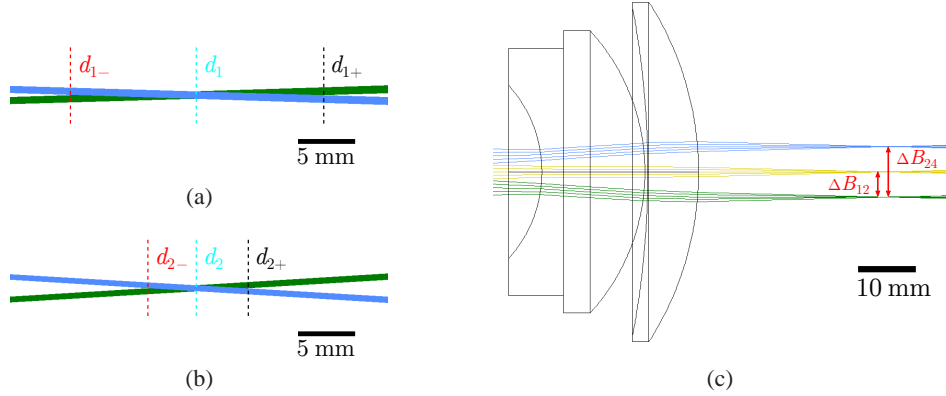


Fig. 7. Zemax screenshots: (a–b) Intersecting light beams at distances $d_{a\pm}$ considering the pixel pitch; (c) Chief rays traveling through Double Gauss objective indicating the baseline ΔB_g .

Figures 7(a) and 7(b) depict screenshots of the distance and depth measurements in Zemax based on the objective lens f_{100} while Fig. 7(c) shows parts of the lens where chief rays reveal baseline results ΔB_{12} and ΔB_{24} in object space.

Table 1. Objective lens parameters.

Objective	Focal length f_U [mm]	Principal plane space $\overline{H_{1U}H_{2U}}$ [mm]
Gauss f_{100}	99.515	-33.184
Nikon f_{200}	193.446	-66.159

Figure 8 shows a plot of the Matlab ray tracing implementation according to the suggestion made in Section 3. Measurements have shown that calculations in Matlab 7.11.0.584 (R2010b) on an Intel Core i7-3770 CPU @ 3.40 GHz require 0.05 s to 0.10 s for the DOF of a single FIS. In contrast, the more laborious design in Zemax includes modelling lenses and measuring distances which can take up to a few hours for an experienced optical engineer.

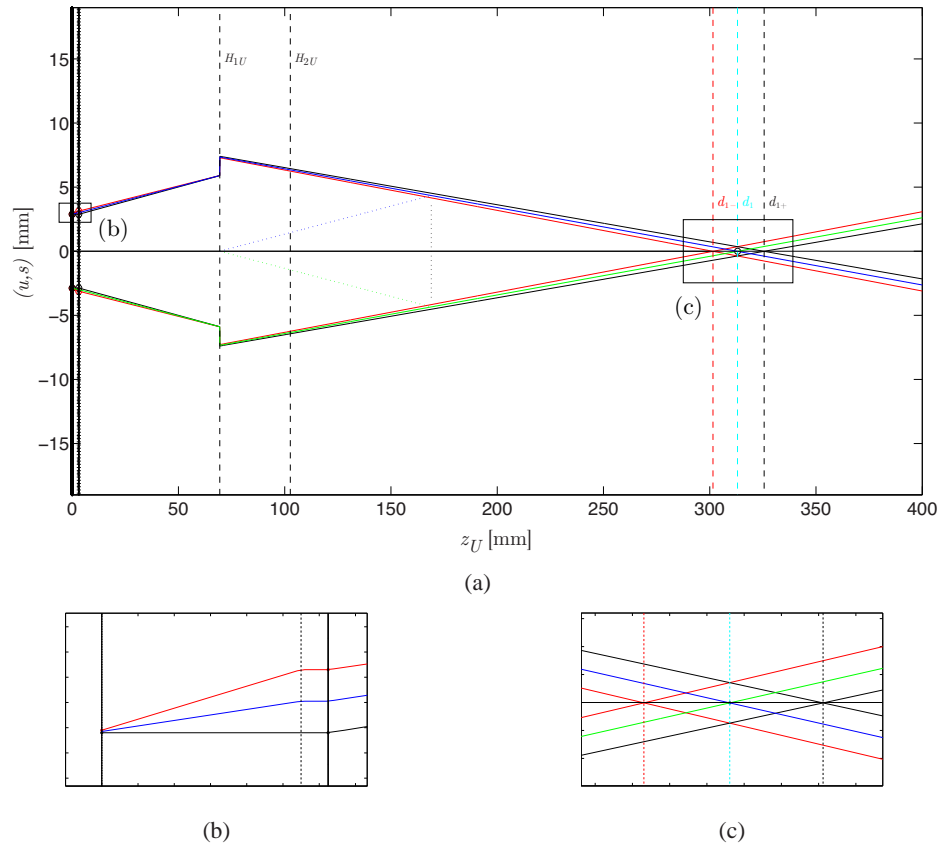


Fig. 8. Matlab screenshots: (a) Paraxial ray tracing based on refraction at principal planes; (b) Close up of rays under a micro lens; (c) Close up of the ray intersection in object space.

When implementing the ray tracing equations, problems arise if the micro image resolution \hat{m} is inconsistent due to the dimensioning of p_p and $p_{\hat{m}}$. One approach to prevent this may be achieved by specifying p_p and $p_{\hat{m}}$ appropriately in advance. Alternatively, one can solve this problem afterwards by resampling the entire raw image, aiming to create a homogeneous micro image resolution. Given the example of having a pixel pitch $p_p = 9 \mu\text{m}$ and a micro lens pitch $p_{\hat{m}} = 250 \mu\text{m}$, the raw image may be downsampled by factor 1.1 yielding $\hat{m} = 25$. In case \hat{m} is even, micro images can be cropped leaving out a pixel in each micro image.

5.2. Results

A data comparison of predicted and simulated distances is shown in Table 2. The deviation ERR of the estimated depth of an FIS is obtained by

$$ERR = \frac{\text{prediction} - \text{simulation}}{\text{prediction}} \times 100. \quad (37)$$

Table 2. Comparison of refocusing distances d_a and $d_{a\pm}$ with respect to the image sensor.

FIS	Prediction [mm]			Simulation [mm]			Deviation [%]		
	d_{a-}	d_a	d_{a+}	d_{a-}	d_a	d_{a+}	ERR_{a-}	ERR_a	ERR_{a+}
0	3483.174	∞	∞	3863.248	∞	∞	-10.912	—	—
1	301.538	313.066	325.596	301.518	312.889	324.189	0.007	0.057	0.432
2	236.606	241.017	245.616	236.565	240.914	245.01	0.017	0.043	0.247
3	214.369	217.001	219.707	214.289	216.809	219.37	0.037	0.088	0.153
4	203.136	204.993	206.889	202.917	204.764	206.647	0.108	0.112	0.117
5	196.359	197.788	199.241	195.99	197.442	198.905	0.188	0.175	0.169

Note that for the respective experiment the separation between the main lens and the MLA has been set to the focal length of the main lens ($b_U = f_U$). Because of that, only *inner rays* intersect at FIS $a = 0$, since *central rays* travel parallel to each other and *outer rays* even diverge. This means that FIS $a = 0$ starts at distance d_{0-} and is all-in-focus from that boundary to infinity. A visualization of the measured values is given in Fig. 9, although d_{0-} has been left out for sake of clarity in the diagram. As seen in the chart, the DOF of an FIS shrinks with increasing a . Interestingly, detailed analysis also indicates that there are uncovered gaps between $d_{a\pm}$. However, these gaps may be retrieved by expanding the numerical range of FISs from $a \in \mathbb{Z}$ to $a \in \mathbb{R}_+$, so that the DOF between integer number FIS is acquired. Referring to Eq. (9), it is apparent that this attempt inherently implies an interpolation of micro images upscaling the spatial resolution at the same time.

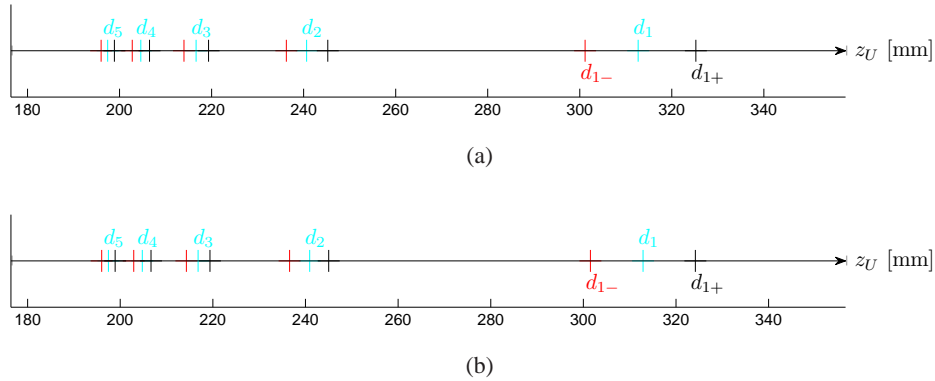


Fig. 9. 1-D plot of (a) predicted and (b) simulated distances d .

Due to a sufficient accuracy of an error of less than 0.5% at $d_{a\pm} \approx 300$ mm, the suggested estimation attempt can contribute to a broader objective such as a depth map generation. Using the results of the proposed distance prediction in combination with a spatial frequency analysis and object segmentation, a novel and competitive depth map method can be developed.

Large errors affect the imaging performance in a way that an object surface, which is supposed to be in focus according to the predictions, may be perceived as slightly blurred. This occurs because the object surface is outside the DOF range, but still very close to its boundary. This observation merely applies to large errors, thus for objects being far away from the capture device. To give an example from Table 2, the paraxial approach predicts an image slice $a=0$ to be in focus from distance $d_{0-} = 3483.174$ mm to infinity, however the FIS actually has its maximal effective resolution from distance $d_{0-} = 3863.248$ mm to infinity. Hence, a predicted synthesized DOF boundary at approximately 3.5 m deviates by 40 cm at the maximum.

Table 3. Refocusable distances d_a with different lens settings for f_s and $b_U \geq f_U$.

FIS	Central distance d_a [mm]						
	f_s	2.749 mm		1.623 mm	2.749 mm		1.623 mm
	b_U	f_{100}	$f_{100} + 10$ mm	f_{100}	f_{200}	$f_{200} + 10$ mm	f_{200}
0		∞	1169.282	∞	∞	4075.980	∞
1		313.066	304.762	412.264	868.363	809.197	1245.355
2		241.017	246.131	290.229	596.109	587.645	784.218
3		217.001	224.779	249.551	505.358	506.962	630.506
4		204.993	213.728	229.212	459.982	465.204	553.650
5		197.788	206.973	217.008	432.757	439.677	507.536

When undertaking a second experiment, focal length parameters f_s , f_U and the lens separation distance b_U have been changed to investigate their impact on the captured LF. Therefore, another micro lens is introduced with the same specification as the previous one apart from the focal length $f_s = 1.623$ mm. Estimated results are listed in Table 3 and depicted in Fig. 10. Table 3 is revealing in several ways. First, it is obvious that closer focusing (in other words shifting the objective lens so that $b_U > f_U$) moves the furthest FIS $a = 0$ towards the camera device. Secondly, decreasing f_s shifts all FISs away from the camera. In addition, it is evident that the larger the value of f_U , the further away the FISs are from the camera and the more space is between them. Hence, it is feasible to shrink and expand the depth of the captured LF which can be useful for photography as it keeps the flexibility to adjust the LF boundaries to landscape or portrait photos, respectively. Furthermore, the distance prediction helps to determine optical parameters of a refocusing-capable SPC in advance.

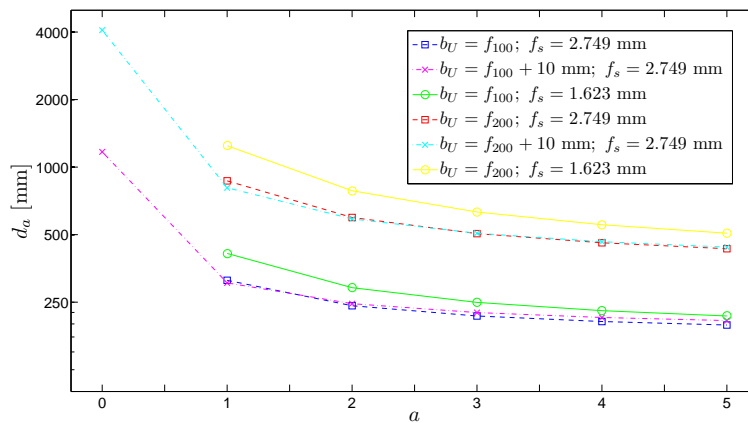


Fig. 10. Impact of varying lens parameters on the refocusing distance d_a .

Conduction of a third experiment aims to validate the baseline estimation presented in Section 4. Given a plenoptic setup composed of the specified MLA with $f_s = 2.749$ mm and the f_{100} objective, virtual lens positions are provided in Table 4. From the data, it is seen that there were no significant differences between the baseline prediction and simulation which proves the suggested concept. An interesting observation is that the baseline ΔB_g does not exceed 10 mm which is relatively small compared to human stereo vision ($\Delta B \approx 65$ mm [19]). This finding confirms the association between plenoptic cameras and microscopy in previous research articles [8,9] as microscopes benefit from the small baseline. However, to apply an SPC in the field of stereo vision, ΔB needs to approach 65 mm. For example, a theoretical baseline $\Delta B_{180} = 65.161$ mm may be accomplished by $p_{\hat{m}} = 2$ mm, $p_p = 0.01$ mm, $f_s = 2.749$ mm and main lens f_{100} producing a decrease in spatial viewpoint image resolution by a factor of 8 in case the sensor and MLA format remain the same. In other words, a large aperture A , a large micro lens pitch $p_{\hat{m}}$ and large sizes in MLA and image sensor will provide larger baselines.

Table 4. Virtual lens positions where $f_U = f_{100}$, $b_{U1} = f_U$ and $b_{U2} = f_U + 20$ mm. The predicted depth position $F_U = 99.514$ mm is given with respect to the principal plane H_{1U} .

Gap	Prediction			Simulation					
	Baseline	b_{U1}	b_{U2}	b_{U1}			b_{U2}		
(i, g)	ΔB_g [mm]	Φ_i [°]	Φ_i [°]	ΔB_g [mm]	F_U [mm]	Φ_i [°]	ΔB_g [mm]	F_U [mm]	Φ_i [°]
(-1, 1)	0.362	0.000	-0.042	0.362	99.485	0.000	0.362	99.491	-0.041
(-12, 12)	4.344	0.000	-0.504	4.341	99.426	-0.001	4.340	99.485	-0.503
(-12, 24)	8.688	0.000	-0.504	8.683	99.426	-0.001	8.681	99.485	-0.503

In the ray tracing intersection model, the SPC is viewed to be a paraxial optical system. Hence, spherical, coma and chromatic aberrations have not been taken into consideration. However, the real ray tracing in Zemax takes into account of aberrations and thus it is assumed that errors represent the deviation due to paraxial approximation disregarding optical aberrations.

6. Conclusions

Apart from distance predictions based on a camera array, to the extent of our knowledge the present work is the very first method providing object space distances of FISs and positions of virtual viewpoint cameras. Given the plenoptic ray tracing model, experimental work using objective lenses allows for the prediction of all available LF slices with a relative error of less than 0.5% with regard to distances of 300 mm and 11% for distances at 3.5 m, whereas baseline estimations of virtual views show no significant difference. Thereby, the suggestion supports to specify the optical parameters. Moreover, the major advantage of the proposed estimations over the complex simulation is an instant computation able to offer several results simultaneously. Implementing such an innovation to an SPC would provide a novel feature. Beyond that, the presented research may be a starting point to develop and evaluate an alternative depth map technique based on distances of the FISs rather than the commonly used disparity analysis.

Acknowledgments

We are grateful to Hector Navarro Fructuoso for the inspiring discussions just as Amy-Grace Douglas and Lascelle Mauricette for the helpful suggestions. This research was supported in part by the University of Bedfordshire and the EU under the ICT program as Project 3D VIVANT under EU-FP7 ICT-2010-248420.